

Übungsblatt 1

Signal Detection Theory (SDT)

In dieser Aufgabe sollen Sie sich praktisch mit den Konzepten der SDT beschäftigen und Erfahrungen in der Einschätzung der Fähigkeit zur Signalerkennung eines Operators sammeln.

Bei dieser Aufgabe ist Gruppenarbeit explizit erwünscht. Bitte bearbeiten und diskutieren Sie alle Aufgaben gemeinsam in Ihrer Gruppe und geben Sie einen **gemeinsamen Bericht** (4-8 Seiten inkl. aller Abbildungen, Tabellen und der Code-Beispiele) ab.

Die Daten

Ihre Gruppe schlüpft in dieser Aufgabe in die Rolle eines mit der Einstellung neuer Mitarbeiter beauftragten Komitees der Personalabteilung an einem Flughafen. Es soll ein Screener eingestellt werden, d.h. ein Beamter, der Bilder von Sicherheitsscannern hinsichtlich deren Bedrohungspotenzials bewertet. Ihre Aufgabe ist es, die Leistung von 200 Bewerbern einzuschätzen und zu versuchen, eine optimale Entscheidung darüber zu treffen, welche Bewerber eingestellt werden sollen.

Während des Bewerbungsgesprächs hat jeder Screener-Bewerber 50 Röntgenbilder von Gepäckstücken angesehen und mit „Stop“ bewertet, wenn eine Bedrohung (z. B. durch Waffen oder Kampfstoffe) festgestellt werden konnte. Ansonsten wurde das Gepäckstück mit „Pass“ durchgewinkt. Die entsprechenden Daten sind in der Datei *sdt-data.csv* enthalten.

Die erste Spalte im Datensatz enthält die Gepäckstücknummer, die zweite Spalte gibt an ob das Gepäckstück tatsächlich eine Waffe enthielt („1“) oder nicht („0“). Jede der nächsten 200 Spalten steht für einen Screener-Bewerber. In diesen Spalten steht eine „1“ für eine **korrekte Bewertung** („Pass“ für Gepäckstücke ohne Waffen, „Stop“ für Gepäckstücke mit Waffen) oder eine „0“ für **falsche Bewertungen** („Pass“ für Gepäckstücke mit Waffen, „Stop“ für Gepäckstücke ohne Waffen). Die untenstehende Tabelle stellt diese Beziehung zwischen Präsenz des Targets in Spalte 2 und den Daten der Screener-Bewerber dar.

		Target Präsenz (@@present)	
		Waffe da=1	0=Keine Waffe
Screener Accuracy (@correct)	1=Korrekte Antwort	Hit	Correct Rejection
	0=Falsche Antwort	Miss	False Alarm

Das Programmgerüst

Um die einzelnen Screener-Bewerber bewerten und eine Entscheidung für die Auswahl der besten Kandidaten treffen zu können, sollen Sie einen automatisierten Klassifikator entwickeln. Dieser soll die Leistung der Bewerber hinsichtlich accuracy, hit rate, miss rate, correct rejection rate, false alarm rate, sensitivity (d') und bias/criterion (C) bewerten.

In der beiliegenden Datei *classifier.py* finden Sie ein Programmgerüst, das bereits in der Lage ist, die Daten einzulesen, für jeden Screener-Bewerber ein kapselndes Objekt anzulegen und nach außen einem Klassifikator die entsprechenden Bewertungsmethoden anbietet. Der Klassifikator enthält Platzhalter für die Implementation eines Rankings, das einzelne oder mehrere Indikatoren verwenden soll, um die Screener-Bewerber in eine geeignete Reihenfolge zu bringen.

Das Programmgerüst ist in der Programmiersprache Python verfasst und enthält eine einfache Klassenhierarchie. Für den Indikator *accuracy* ist bereits eine Implementierung vorhanden, die restlichen Indikatoren sind vorgesehen, geben bislang aber immer 0 zurück. Für Ihre Modifikationen des Programmgerüsts sollten nur Änderungen zwischen den # TODO #-Blöcken nötig sein.

Bericht

Bitte strukturieren Sie Ihre Abgabe anhand der folgenden Vorgabe.

Deckblatt: mit Namen aller Gruppenmitglieder und deren Matrikelnummern, Übungsnummer (z.B. Übungsblatt 1) und Titel der Aufgabe.

1. Einleitung

- Was ist die SDT? Geben Sie eine kurze Beschreibung.
- Geben Sie eine kurze Definition der folgenden Begriffe an: hit, miss, false alarm, correct rejection, sensitivity, bias und response criterion
- Was ist für Sie der wichtigste Indikator bei der Auswahl eines Screeners? Warum?

2. Methode/Ergebnisse

- a. Beschreiben Sie die Implementation der false alarm rate ausführlich!
- b. Fügen Sie dem Programmgerüst eine Methode hinzu, um die durchschnittliche hit rate, miss rate, false alarm rate und correct rejection rate für die gesamte Gruppe von Screener-Bewerbern zu ermitteln. Stellen Sie die ermittelten Daten in einer Matrix dar (wie in der Übung besprochen).
- c. Entwickeln Sie die Classifier-Klasse so weiter, dass die Leistung der Screener-Bewerber anhand der Indikatoren bewertet wird und eine Reihung erstellt wird. Setzen Sie verschiedene Varianten um:
 - i. Klassifikation nur anhand der accuracy
 - ii. Klassifikation nur anhand der hit rate
 - iii. Klassifikation nur anhand der false alarm rate
 - iv. Kombinierte Klassifikation mit 2 Indikatoren

3. Diskussion: bitte beantworten/diskutieren Sie die folgenden Bereiche

- a. Schauen Sie sich die Ergebnisse des Klassifikators aus Aufgabe 2c an und entscheiden Sie pro Klassifikation (accuracy, hit rate, false alarm rate, zwei Indikatoren) welchen der Top-5-Bewerber Sie einstellen würden. Warum?
- b. Sensitivity (d')
 - i. Welcher Bewerber hat die höchste sensitivity, welcher die geringste?
 - ii. Haben 2 der Bewerber ungefähr dieselbe sensitivity, aber unterschiedliche hit- und false alarm rates? Was könnten Sie tun, um 2 Screener zu finden, die sich bei gleicher sensitivity ähnlicher sind?
- c. Bias/criterion (C)
 - i. Welche 2 Bewerber hatten den liberalsten Bias? Welche 2 hatten den konservativsten Bias? Beschreiben Sie kurz: Was bedeutet das?
 - ii. Welche Ereignisse könnten im Kontext des Screenings für Flughafensicherheit dazu führen, dass ein Screener einen liberaleren Bias erlangt? Was könnte zu einem konservativeren Bias führen? Diskutieren Sie sowohl den Erwartungswert als auch Kosten/Nutzen.

4. Fazit

- a. Auf der Grundlage **aller** zuvor ermittelten Daten: wie lautet Ihre finale Einstellungsentscheidung? Hat sich der anfänglich gewählte Indikator zur Auswahl des Screeners geändert? Bitte geben Sie auch ihre 4 weiteren Top-Favoriten für die Stelle an und begründen Sie ihre Wahl kurz.
- b. Was würden Sie hinsichtlich Arbeitsumgebung, Arbeitszeiten, Ausbildung oder sonstigem im Bereich des Screenings an Flughäfen verändern, um die Leistung der Screener weiter zu optimieren?

Anhang: Bitte fügen Sie ihrer Abgabe den finalen Quelltext als .py-Datei hinzu.

Bewertungsübersicht

		Maximale Punkte		Erreichte Punkte	
1. Einleitung	a) Was ist SDT?	3	8,5		
	b) Definitionen der Indikatoren	3,5			
	c) Indikatoren	2			
2. Methode/ Ergebnisse	a) Beschreibung der Indikatoren	2	12,5		
	b) Gruppenergebnisse	8,5			
	c) Klassifikationen	2			
3. Diskussion	a) Einstellungsentscheidung	4	14		
	b) Sensitivity	4			
	c) Bias	6			
4. Fazit	a) Finale Einstellungsentscheidung	5	10		
	b) Möglichkeiten der Leistungsoptimierung	5			
<i>Qualität insgesamt: Tiefe der Diskussion, Anschaulichkeit, Grammatik, Rechtschreibung, Formatierung</i>		5	5		
Gesamt		50			

Die Abgabe der Lösung (PDF/ZIP) muss bis Sonntag, 23:59 Uhr in 2 Wochen (23.04.2017) erfolgen! Verspätete Abgaben erhalten einen Punktabzug!